

SEQUENCE LISTING

SEQ ID NOs:1 and 2

LOCUS COTMYBA 1006 bp mRNA PLN 31-DEC-1993

DEFINITION Cotton DNA-binding domain mRNA.

ACCESSION L04497

NID g437326

VERSION L04497.1 GI:437326

KEYWORDS .

SOURCE Gossypium hirsutum (cultivar Acala SJ-2) 3-day pre-anthesis ovule

cDNA to mRNA.

ORGANISM Gossypium hirsutum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 1006)

AUTHORS Wilkins,T.A. and Lu,C.-C.

JOURNAL Unpublished (1993)

FEATURES Location/Qualifiers

source 1..1006

/organism="Gossypium hirsutum"

/cultivar="Acala SJ-2"

/db_xref="taxon:3635"

/dev_stage="3-day pre-anthesis"

/tissue_type="ovule"

mRNA 1..1006

CDS 59..943

/note="MYB A; putative"

/codon_start=1

/protein_id="AAA33067.1"

/db_xref="PID:g437327"

/db_xref="GI:437327"

protein_bind 92..403

/note="putative"

/bound_moiety="MYB"

repeat_region 107..133

/note="MYB DNA-binding domain repeat signature 1;
putative"

misc_feature 323..394

BASE COUNT 323 a 209 c 204 g 270 t

ORIGIN

1 taacaccgtt attctttctc tattctacct gatttgattt gatttgattt tgtaactgat

61 gggacgatca ccttggttg aaaaggctca taccaacaaa ggtgcctgga ccaaagagga

121 agatcaacgc ctcacaaact acatccgtgt ccatggtgaa ggctgctggc gttccctccc

181 caaagctgct gggctgctta gatgtggtaa gagttgcaga ttaagatgga taaactactt

241 gaggcctgat cttaagagag gaaatttcac tgaagaagaa gatgagctta tcacaaagct

301 tcacagttaa cttggaacaa aatgggtcatt gattgctgga agattaccag gaagaacaga

361 taatgagata aagaactact ggaacacaca catcaaaaga aagcttataa gcagaggaat

421 tgatccacaa actcatcgtc ctctcaatca aacggccaat accaacacag tcacagcccc

481 caccgaattg gatttcagaa actcgccac atccgtttcc aaatccagtt ccatcaaaaa
 541 cccgtctctg gatttcaatt acaatgaatt tcaattcaag tccaacacag attcccttga
 601 agaaccacac tgtacagcca gcagtggcat gactacagat gaagagcaac aagaacagct
 661 gcacaagaag cagcaatacg gtccgagcaa tgggcaagac ataaatttgg agctgtcgat
 5 721 tgggattgtt tcagctgact catctcgggt atcaaatgcc aactcggccg agtcgaaacc
 781 aaaggttagat aacaacaatt tccagtttct tgaacaagct atggtggcta aggcggtatg
 841 tttgtgttgg caattaggtt ttggaacaag tgaatttgt aggaactgtc aaaattcaaa
 901 ttcaaatggc ttctatagtt atttagacc cttggattca tagggtcatc tttttctct
 961 ttctttctgt ttttaggaga taaattaatg ctttaattatt aaaaaa

10

MGRSPCCEKAHTNKGAWTKEEDQRLINYIRVHGEGCWRSPLKAALLRCGKSCRLR
 WINYLRPDLKRGNFTEEEDELIKLHSLGKWSLIAGRLPGRDTDNEIKNYWNTHIKR
 KLISRGIDPQTHRPLNQTANTNTVTAPTELDNRNSPTSVSKSSSIKNPSLDFNYNEFQF
 15 KSNTDSLEEPNCTASSGMTTDEEQQLHKKQQYGPSNGQDINLELSIGIVSADSSRV
 SNANSAESKPKVDNNNFQFLEQAMVAKAVCLCWQLGFGTSEICRNCQNSNSNGFY
 YCRPLDS

20 **SEQ ID NO:3 and 4**

LOCUS AF034134 1151 bp mRNA PLN 02-MAR-1998
 DEFINITION Gossypium hirsutum MYB-like DNA-binding domain protein (Cmy-O)
 mRNA, complete cds.

ACCESSION AF034134
 25 NID g2921339
 VERSION AF034134.1 GI:2921339
 KEYWORDS

SOURCE upland cotton.
 ORGANISM Gossypium hirsutum
 30 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 1151)
 AUTHORS Loguercio,L.L., Zhang,J. and Wilkins,T.A.
 35 TITLE Structure and expression of six classes of myb-domain genes in
 allotetraploid cotton (Gossypium hirsutum L.)
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1151)
 AUTHORS Loguercio,L.L., Zhang,J. and Wilkins,T.A.
 40 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1997) Agronomy & Range Science, University of
 California, One Shields Ave., Davis, CA 95616-8515, USA

FEATURES Location/Qualifiers
 source 1..1151

45 /organism="Gossypium hirsutum"
 /cultivar="Acala SJ-2"
 /db_xref="taxon:3635"
 /dev_stage="3 days pre-anthesis"
 /tissue_type="ovule"

gene 1..1151
 /gene="Cmy-O"
 /note="MYB-domain gene O"
 CDS 72..752
 5 /gene="Cmy-O"
 /function="putative MYB-like transcription factor"
 /note="similar to MYB A encoded by GenBank Accession
 Number L04497"
 /codon_start=1
 10 /product="MYB-like DNA-binding domain protein"
 /protein_id="AAC04720.1"
 /db_xref="PID:g2921340"
 /db_xref="GI:2921340"

15 BASE COUNT 382 a 215 c 240 g 314 t
 ORIGIN

1 cggattttct tccccgtgt ttggtgcac agaaagttag agaaagtttt acttttgatt
 61 ttgaaactcc gatgagaaaa ccttgctgcg ataaacaagg caccaacaag ggagcctggt
 121 ccaagcaaga agatcaaaaag ctcattgatt atatacgtat tcatggtgaa ggctgttgcc
 20 181 gtccctccc caaagctgca ggtttgcacc gttgcggtaa aagttgcagg ctgagatgga
 241 taaattactt aagaccagat atcaaactg gtaacttgc tcaagacgaa gaggacttaa
 301 ttatcaactt ccatgctctc ctggtaacc ggtggtcact gatagctggt agattaccag
 361 gaagaacaga taatgaagtg aagaactatt ggaattccca tataaagaga aagctaata
 421 agatggggat cgatectaata aaccataagt tgaaccaata tctcatcat gttggtcccc
 25 481 ttaacccac caccaccaac tccatggatg tggcatgtaa gcttagagtg tgtcaacag
 541 acaatgatga tgggatctca gatgctgcaa gttatctga agacgcaaca ccgcccactg
 601 gtatatccaa ctggacctt gatctcacia ttgctttcc ttcgagtcct atcaagaata
 661 ttattgaaga aagccagcag aaaacagcat ctattgtaac aaatgatgaa gaagaacaat
 721 atacagtccc tacccttctt ctttcagat gagacaaaaa aaaaagcctc acacatgtgg
 30 781 agattcgtgc aaaagaccta aaggcttacg aaggcaacat gcacgccatt gtcaaattct
 841 ttgggatgat ggattgaac catatccttg tccattagaa aggaggaaga taagctaaaa
 901 ctgtattatt gtgtataaat ttggtagaaa gaaagatttc aacttaagaa ttaggatcaa
 961 ataactgaat gaatgaacga attgcagata agttgttagg aggtttcaa tcaacttatc
 1021 tgcaattaat ttggtggagc tgatgtagga tgatgagttc atcgtacatg aactgaacct
 35 1081 ttgatatttc aggctctaata tgtctgttg tatgcgtaaa gatattcttc aatgtgagat
 1141 cagctaaaaa a

MRKPCCDKQGTNKGAWSKQEDQKLIDYIRIHGEGCWRSPLKAAGLHRCGKSCRLR
 WNYLRLPDIKRGNFQAQDEEDLIKLHALLGNRWSLIAGRLPGRTDNEVKNYWNSHIK
 40 RKLMKMGIDPNNHKLNLQYPHHVGPLNPTTTNSMDVACKLRVCSTDNDDGISDAAS
 YLEDATPPTGISNLDLDTIAFPSSPIKNIIIESQQKTASIVTNDEEEQYTVPTLLLFR
 repeat_region 102..257

/note="putative MYB DNA-binding domain repeat R2"

repeat_region 258..412

/note="putative MYB DNA-binding domain repeat R3"

SEQ ID NO: 5

Cotton (*Gossypium hirsutum* L. cv. Acala SJ-2) 10 dpa fiber cDNA

R2R3-MYB transcription factor

GhMYB7

ID 6W DNA; 1081 BP.

SQ SEQUENCE 1081 BP; 353 A; 228 C; 240 G; 260 T;

CTCCCCGCGG TGGCGGCCGC TCTAGAACTA GTGGATCCCC CGGGCTGCAG
GAATTCGGCA

CGAGGAAAGA AGTGTGAAAA AAAAAATGGG AAGGAGTCCT TGTTGTTCTA
AGGAAGGCCT

TAACAGAGGA GCTTGGACTG CTCTTGAAGA CAAAATTCTT AAAGATTATA
TCAAAGTACA

CGGTGAAGGT CGTTGGAGAA ATCTCCCCAA AAGAGCTGGT CTTAAGAGAT
GTGGGAAAAG

TTGTAGGCTT CGGTGGTTGA ATTATTTGAG ACCTGATATT AAAAGAGGTA

ACATATCACC

TGACGAGGAA GAGCTTATCA TCAAACCTCA CAAACTCTTG GGAAACAGAT
GGTCTTTGAT

AGCTGGGAGG CTTCCAGGAC GAACAGACAA TGAAATAAAG AATTACTGGA
ACACCAACTT

AAGTAAAAGA GTTTCCGATC GTCAAAAGTC ACCCGCCGCT CCTTCGAAAA
ATCCCGAGGC

GGCTCGACGA GGAAGTCTG GTAATGGCAA TACCAATGGT AATGGTAGTG
GTAGTTCCTC

GACACACGTG GTGCGGACAA GGGCGACAAG GTGCTCCAAG GTTTTCATAA
ACCCTCCTCA

CTACACACAA AACAGAGACC CAAAGCCTTC TTCAACTTGT TCAAATCATG
GGGATCACCG

GGAACCTAAA ACAATGAATG AGTTGTTATT ACCGATAATG TCAGAATCCG
AGAATGAAGG

GACGACCGAT CATATATCAT CGGATTTTAC ATTTGACTTC AACATGGGAG
AGTTTTGTTT

ATCGGATCTT TTGAATTCCG ATTTCTGCGA TGTAACGAG CTTAATTACA
GCAATGGTTT

TGATTCGTCA CCCTCACCGG ATCAGCCTCC TATGGATTTC TCCGACGAAA
TGCTAAAAGA

GTGGACGGCC GCCGCCTCCA CTCACTGCTG TCACCAAAGT GCGGCTTCCA
ATCTCCAGTC

CTTGCCTCCA TTTATTGAAA ATGGAATTGA ATGACCTTGA AAAAATAAAA
GACGAAAAAT

ATTTTCTCAT GTAAACTAAA TAAACACATC TTCCATCATT AAAAAAAAAA
AAAAAAAAAA

A

/

SEQ ID NO:6

ID 6WP PRT; 302 AA.

DT 09-NOV-1999 (CREATED BY PC/GENE PROGRAM TRANSL)

5 DE

CC TRANSLATED FROM DNA SEQUENCE 6W (BASES 86 TO 993).

SQ SEQUENCE 302 AA; 33766 MW; 451675 CN;

MGRSPCCSKE GLNRGAWTAL EDKILKDYIK VHGEGRWRNL PKRAGLKRCG
KSCRLRWLNY

10 LRPDIKRGNI SPDEEELIHK LHKLLGNRWS LIAGRLPGRT DNEIKNYWNT
NLSKRVSDRQ

KSPAAPSKNP EAARRGTAGN GNTNGNGSGS SSTHVVRTRA TRCSKVFINP
PHYTQNRDPK

15 PSSTCSNHGD HREPKTMNEL LLPIMSESEN EGTTDHISD FTFDFNMGEF
CLSDLLNSDF

CDVNELNYSN GFDSSPSPDQ PPMDFSDEML KEWTAAASTH CCHQSAASNL
QSLPPFIENG

IE

20

GeneBank

Cotton (*Gossypium hirsutum* L. cv Acala SJ-2) 10 dpa fiber cDNA
R2R3-MYB Transcription Factor
GhMYB8

5 **SEQ ID NO: 7**

ID 19W; DNA; 944 BP.

SQ SEQUENCE 944 BP; 319 A; 163 C; 214 G; 246 T; 2 OTHER;

TGGAGCTCCC CGCGGTGGCN GNCGCTCTAG AACTAGTGGA TCCCCCGGGC
TGCAGGAATT

10 CGGCACGAGA CTCCAACAAA TGTCCATGAA AAAAGAAGGT GAAATTCTAT
ACAAAAAGGG

ATTATGGGCA ATGGAGGAAG ACAAGTTACT CATTGATTAT GTCAATGTCC
ATGGAAAAGG

15 ACAATGGAAC AAAATAGCCA ACAGAACAGG TTTGAAGAGA AGTGGGAAAA
GTTGTCGGCT

AAGGTGGATG AATTACCTGA GTCCTAACGT TAAAAAGGGT GATTTTTCTG
AAGAAGAAGA

AGACCTCGTC ATTAGACTTC ATAAGCTTCT TGGAAACAGG TGGTCTTTGA
TTGCGAAACG

20 AGTTCCAGGT CGAACTGACA ATCAAGTCAA GAATTACTGG AATAGTCATT
TGAGGAAGAA

ACTAGGGATC ATTGATCAAA ACAAGACAAG GATCGATTTT TGTCAAAGTT
CAAAGCAAGT

25 CAAAGTGTGT CATGTTGATG AGGCAGCCAC GGATCCAAGT CCTGGACATG
GAACAACCAC

TGAAACCACG GGTATAACAG TGGATCAGAG TAACCAGCAG GAAGTCATTG
ATCATCGGGT

CTTAAACAAT ACTACTCAAG AATCAATGAC CAGTGAGAGT TATATCAACA
CTTTCTGGAT

30 TCCTGACCAT GATTATGAGC TAAGTACACT TGCCATGATT GACCATGATT
ATGAGCTAAG

TACACTTGCC ATGATTGACC ACTTCCATGA ATGTTCTTCT TTTCATCTTA
GCTAGAGACT

35 ATGTTATTAG ATTCGGGTTT TATTTTTAGA TATAAGTATG CATCTAACAT
GGCAATGTTA

AATTTTTCAA AAGATTTTTTC ATGTATTTGA GCAGTTCATG TGTTTGAAGA
TTAAGATATT

TCTGAAAAAA AAAAAAAAAA AAAACCGAGG GGGGCCCGGT ACCC

//

40

SEQ ID NO:8

ID 19WP; PRT; 231 AA.

DT 25-OCT-1999 (CREATED BY PC/GENE PROGRAM TRANSL)

45 DE

CC TRANSLATED FROM DNA SEQUENCE 19W (BASES 80 TO 772).

SQ SEQUENCE 231 AA; 26698 MW; 283450 CN;

MSMKKEGEIL YKKGLWAMEE-DKLLIDYVNV-HGKGQWNKIA-NRTGLKRSRK
SCRLRWMNYL

NKTRIDFCQS SKQVKVCHVD EAATDPSPGH GTTETTGIT VDQSNQQEVI
DHRVLNNTTQ

5